

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
**TRƯỜNG ĐẠI HỌC ĐẠI NAM**  
**KHOA CÔNG NGHỆ THÔNG TIN**

---



**BÁO CÁO BÀI TẬP LỚN**

**Ứng dụng giải thuật di truyền trong tối ưu sắp xếp chỗ ngồi  
tiệc cưới**

**Sinh viên thực hiện : Nguyễn Quốc Việt  
Nguyễn Thị Phương  
Kiều Văn Phúc  
Nguyễn Tiến Đạt**

**Ngành : Công nghệ thông tin  
Giảng viên hướng dẫn : ThS. Lê Thị Thùy Trang**

## Lời cảm ơn

Công trình này có sử dụng một phần sản phẩm của và **Ngo Thi Tien**.

### **Tóm tắt**

Tài liệu này cung cấp một bản thiết kế mẫu biên soạn báo cáo bài tập lớn theo quy định của Trường đại học Đại Nam.

# Mục lục

<b>1</b>	<b>Giới thiệu</b>	<b>1</b>
1.1	Giới thiệu tổng quan về đề tài . . . . .	1
1.2	Mục đích và lý do lựa chọn đề tài . . . . .	2
<b>2</b>	<b>Cơ sở lý thuyết</b>	<b>4</b>
2.1	Giới thiệu về giải thuật di truyền . . . . .	4
2.2	Các thành phần chính . . . . .	11
2.3	Ứng dụng giải thuật di truyền trong các bài toán tối ưu . . . . .	14
<b>3</b>	<b>THIẾT KẾ VÀ GIẢI THUẬT</b>	<b>18</b>
3.1	Mô hình bài toán . . . . .	18
3.2	Mã hóa cá thể . . . . .	19
3.3	Hàm thích nghi (Fitness Function) . . . . .	20
3.4	Toán tử lai ghép và đột biến . . . . .	20
3.4.1	Toán tử đột biến (Mutation) . . . . .	21
3.4.2	Phương pháp . . . . .	21
3.5	Chiến lược chọn lọc . . . . .	21
3.5.1	Mục đích . . . . .	21
3.5.2	Phương pháp . . . . .	21
3.5.3	Quy trình: . . . . .	22
<b>4</b>	<b>Kết Quả</b>	<b>23</b>
4.1	Thiết lập dữ liệu . . . . .	23
4.1.1	Môi trường thực hiện . . . . .	23
4.1.2	Dữ liệu đầu vào . . . . .	23
4.1.3	Ràng buộc . . . . .	24
4.1.4	Thông số GA . . . . .	24
4.1.5	Phương pháp chạy . . . . .	24
4.2	Kết quả thu được . . . . .	25
4.3	Đánh giá hiệu suất của giải thuật . . . . .	25

4.3.1	Khả năng đáp ứng ràng buộc . . . . .	25
4.3.2	Thời gian thực thi . . . . .	25
4.3.3	Ưu điểm . . . . .	25
4.3.4	Hạn chế . . . . .	25
4.3.5	Đề xuất cải tiến . . . . .	26
<b>5</b>	<b>Kết Luận</b>	<b>27</b>
5.1	Điểm mạnh của nghiên cứu . . . . .	27
5.1.1	Tính hiệu quả của giải thuật di truyền . . . . .	27
5.1.2	Tính linh hoạt và khả năng mở rộng . . . . .	28
5.1.3	Ứng dụng công nghệ vào đời sống thực tế . . . . .	28
5.2	Hạn chế và thách thức . . . . .	28
5.2.1	Độ phức tạp tính toán và thời gian xử lý . . . . .	28
5.2.2	Khả năng đảm bảo nghiệm tối ưu tuyệt đối . . . . .	28
5.3	Tính chủ quan trong đánh giá mức độ hài lòng . . . . .	28
5.3.1	Tính ứng dụng thực tế và hướng phát triển tương lai . . . . .	29

# **Danh sách bảng**

# Danh sách hình vẽ

4.1	Kết quả chạy được . . . . .	24
-----	-----------------------------	----

# **Danh sách giải thuật**



# Chương 1

## Giới thiệu

### 1.1 Giới thiệu tổng quan về đề tài

Trong các sự kiện quan trọng như đám cưới, việc sắp xếp chỗ ngồi cho khách mời không chỉ đơn thuần là bố trí vị trí ngồi, mà còn ảnh hưởng trực tiếp đến trải nghiệm và sự hài lòng của khách mời. Một sơ đồ chỗ ngồi hợp lý có thể tạo ra không khí thân thiện, giúp gắn kết các mối quan hệ và hạn chế những tình huống không mong muốn. Tuy nhiên, việc sắp xếp này trở nên phức tạp khi số lượng khách mời lớn, có nhiều mối quan hệ chồng chéo và những yêu cầu đặc biệt từ các bên liên quan.

Giải thuật di truyền (Genetic Algorithm - GA) là một phương pháp tính toán dựa trên cơ chế tiến hóa và di truyền trong tự nhiên, thường được áp dụng để giải quyết các bài toán tối ưu phức tạp. Với bài toán sắp xếp chỗ ngồi trong đám cưới, giải thuật di truyền cung cấp một hướng tiếp cận linh hoạt và hiệu quả, cho phép tìm kiếm các phương án sắp xếp tối ưu theo những tiêu chí được đề ra, chẳng hạn như việc đảm bảo khách quen biết nhau ngồi gần nhau, tránh xếp chỗ những người có mối quan hệ không tốt cạnh nhau và tối ưu hóa không gian tổ chức.

Nghiên cứu này tập trung vào việc áp dụng giải thuật di truyền nhằm xây dựng một mô hình tự động hóa quy trình sắp xếp chỗ ngồi cho khách mời trong các sự kiện cưới hỏi. Bằng cách mô phỏng cơ chế chọn lọc tự nhiên, lai ghép và đột biến, giải thuật di truyền có thể tạo ra nhiều phương án sắp xếp và chọn ra phương án tốt nhất, giúp giảm thiểu thời gian và công sức cho việc tổ chức.

Việc áp dụng giải thuật di truyền vào bài toán này không chỉ mang lại tính ứng dụng cao trong thực tế, mà còn thể hiện sự hiệu quả của các phương pháp tính toán thông minh trong việc giải quyết những bài toán mang tính tối ưu hóa tổ hợp phức tạp.

## 1.2 Mục đích và lý do lựa chọn đề tài

Trong các sự kiện quan trọng như đám cưới, việc sắp xếp chỗ ngồi cho khách mời không chỉ là một công việc tổ chức đơn thuần, mà còn ảnh hưởng trực tiếp đến bầu không khí của buổi tiệc và trải nghiệm của các khách mời. Một sơ đồ chỗ ngồi hợp lý sẽ giúp tạo ra sự kết nối, xây dựng các mối quan hệ mới và hạn chế những tình huống không mong muốn do mâu thuẫn cá nhân. Tuy nhiên, việc tìm ra phương án sắp xếp tối ưu trở nên vô cùng phức tạp khi số lượng khách mời lớn, có nhiều nhóm quan hệ đan xen và những yêu cầu đặc biệt từ các bên liên quan.

Mục đích chính của đề tài này là xây dựng một mô hình tự động hóa quy trình sắp xếp chỗ ngồi cho khách mời trong đám cưới bằng cách áp dụng giải thuật di truyền (Genetic Algorithm - GA). Thông qua việc áp dụng giải thuật này, đề tài hướng tới việc cải thiện hiệu quả trong công tác tổ chức, tối ưu hóa quá trình sắp xếp, đồng thời đảm bảo các tiêu chí như: tạo ra không gian giao tiếp thoải mái giữa các khách mời, hạn chế xung đột và tối ưu hóa diện tích sử dụng.

Lý do lựa chọn giải thuật di truyền xuất phát từ những đặc tính vượt trội của phương pháp này trong việc giải quyết các bài toán tối ưu tổ hợp phức tạp. Cụ thể:

- Giải quyết bài toán tối ưu nhiều ràng buộc: Bài toán sắp xếp chỗ ngồi có rất nhiều yếu tố cần cân nhắc như mối quan hệ giữa khách mời, số lượng bàn, sức chứa của mỗi bàn, sở thích cá nhân, và thậm chí là những mâu thuẫn tiềm ẩn giữa một số khách mời. Giải thuật di truyền, với cơ chế mô phỏng quá trình tiến hóa tự nhiên thông qua các phép toán chọn lọc, lai ghép và đột biến, cho phép tạo ra nhiều phương án sắp xếp đa dạng và tìm ra phương án tốt nhất qua từng thế hệ..
- Tính linh hoạt và khả năng tìm kiếm giải pháp tối ưu: Trong khi các phương pháp tìm kiếm truyền thống có thể dễ dàng mắc kẹt ở các nghiệm cục bộ, giải thuật di truyền có khả năng mở rộng không gian tìm kiếm và liên tục cải thiện chất lượng nghiệm qua từng thế hệ. Điều này đặc biệt phù hợp với bài toán sắp xếp chỗ ngồi, nơi tồn tại rất nhiều phương án khả thi và việc tìm ra phương án tối ưu là một thử thách lớn.
- Tăng cường tự động hóa và giảm thiểu sai sót: Quy trình sắp xếp chỗ ngồi truyền thống thường dựa vào kinh nghiệm của người tổ chức, dễ dẫn đến sai sót hoặc mất nhiều thời gian khi quy mô khách mời lớn. Việc áp dụng giải thuật di truyền giúp tự động hóa quá trình này, giảm thiểu đáng kể công sức của con người, đồng thời đảm bảo tính khách quan và khả năng tìm ra các phương án sắp xếp tối ưu trong thời gian ngắn.
- Ứng dụng thực tiễn cao: Đề tài không chỉ dừng lại ở việc giải quyết bài toán sắp xếp chỗ ngồi cho đám cưới, mà còn mở ra tiềm năng ứng dụng rộng rãi trong các sự kiện khác như hội nghị, tiệc công ty hay thậm chí là việc sắp xếp chỗ ngồi trong lớp học, rạp chiếu phim và nhiều bối cảnh khác. Việc nghiên cứu và triển khai thành công mô hình này sẽ tạo tiền

đề cho những ứng dụng thực tiễn, góp phần nâng cao hiệu quả trong công tác tổ chức sự kiện.

# Chương 2

## Cơ sở lý thuyết

### 2.1 Giới thiệu về giải thuật di truyền

#### 2.1.1. Khái niệm

Giải thuật di truyền (Genetic Algorithm – GA) là một phương pháp tìm kiếm và tối ưu hóa dựa trên cơ chế chọn lọc tự nhiên và di truyền sinh học, lấy cảm hứng từ thuyết tiến hóa của Charles Darwin. Phương pháp này mô phỏng quá trình tiến hóa trong tự nhiên, nơi các cá thể ưu tú có khả năng sinh tồn và truyền đặc điểm tốt cho thế hệ sau, dần dần làm quần thể tiến hóa theo hướng tối ưu hơn.

Trong GA, mỗi lời giải tiềm năng cho bài toán được mã hóa thành một cá thể (*individual*), thường biểu diễn dưới dạng một chuỗi gen (*genotype*). Các cá thể này tập hợp lại thành một quần thể (*population*), đại diện cho tập hợp các lời giải tại mỗi thế hệ. Mỗi cá thể được đánh giá thông qua một hàm thích nghi (*fitness function*), thể hiện mức độ phù hợp của lời giải đó đối với bài toán cần giải quyết. Quá trình tiến hóa trong GA diễn ra qua nhiều thế hệ, với ba toán tử chính:

- **Lựa chọn (Selection):** Chọn ra các cá thể ưu tú từ quần thể dựa trên giá trị hàm thích nghi. Các cá thể có giá trị thích nghi cao hơn sẽ có cơ hội được chọn nhiều hơn, mô phỏng quá trình “sinh tồn của kẻ thích nghi nhất”. Một số phương pháp lựa chọn phổ biến là lựa chọn tỉ lệ theo độ thích nghi (*Roulette Wheel Selection*), lựa chọn theo giải đấu (*Tournament Selection*), và lựa chọn theo xếp hạng (*Rank Selection*).
- **Lai ghép (Crossover):** Kết hợp thông tin di truyền của hai cá thể được chọn để tạo ra cá thể con mới, tương tự như quá trình giao phối trong tự nhiên. Phương pháp lai ghép phổ biến gồm lai một điểm (*Single Point Crossover*), lai hai điểm (*Two Point Crossover*), và lai đồng nhất (*Uniform Crossover*).

- **Đột biến (Mutation):** Thay đổi ngẫu nhiên một hoặc một vài gen trong cá thể con để tăng tính đa dạng của quần thể, tránh rơi vào cực trị địa phương. Các phương pháp đột biến thường gặp là đột biến nhị phân (*Binary Mutation*) và đột biến hoán vị (*Swap Mutation*).

Chu trình này được lặp lại cho đến khi đạt được tiêu chí dừng, chẳng hạn như khi số thế hệ đạt ngưỡng tối đa hoặc khi tìm được lời giải tối ưu.

Giải thuật di truyền có ưu điểm là khả năng tìm kiếm toàn cục mạnh mẽ và linh hoạt, đặc biệt hiệu quả trong các bài toán tối ưu hóa phi tuyến tính, không liên tục và không khả vi. Tuy nhiên, GA cũng có nhược điểm là chi phí tính toán lớn do phải xử lý nhiều cá thể trong quần thể qua nhiều thế hệ.

Với sự linh hoạt và khả năng giải quyết các bài toán có không gian tìm kiếm phức tạp, giải thuật di truyền được áp dụng rộng rãi trong nhiều lĩnh vực như: tối ưu hóa tổ hợp (*Combinatorial Optimization*), học máy (*Machine Learning*), lập lịch (*Scheduling*), thiết kế mạng nơ-ron nhân tạo, và các bài toán thực tế như điều phối nguồn lực, thiết kế hệ thống và mô phỏng sinh học.

### 2.1.2. Nguyên lý hoạt động

Giải thuật di truyền (Genetic Algorithm – GA) mô phỏng quá trình tiến hóa trong tự nhiên, nơi các cá thể tốt nhất có cơ hội sinh tồn cao hơn và truyền lại đặc tính ưu việt cho thế hệ sau. Nguyên lý hoạt động của GA bao gồm các bước chính như sau:

#### 1. Khởi tạo quần thể (Initialization):

- **Mục tiêu:** Tạo ra một tập hợp ban đầu gồm nhiều cá thể, mỗi cá thể là một lời giải tiềm năng cho bài toán.
- **Thực hiện:**
  - Xác định kích thước quần thể  $NNN$  (thường từ vài chục đến vài trăm cá thể).
  - Mã hóa các cá thể:
    - \* *Mã nhị phân:* Biểu diễn cá thể bằng chuỗi các bit 0 và 1.
    - \* *Mã hóa thực:* Sử dụng các giá trị thực cho bài toán tối ưu liên tục.
    - \* *Mã hoán vị:* Dùng trong các bài toán hoán vị, như sắp xếp chỗ ngồi, lập lịch,...
  - Các gen trong mỗi cá thể được khởi tạo ngẫu nhiên hoặc dựa trên một số quy tắc nhất định.

#### 2. Đánh giá độ thích nghi (Evaluation):

- **Mục tiêu:** Đo lường chất lượng của từng cá thể trong quần thể.
- **Thực hiện:**
  - Xây dựng một hàm thích nghi (*Fitness Function*) để đánh giá mức độ phù hợp của mỗi cá thể với mục tiêu tối ưu hóa.
  - Giá trị thích nghi thể hiện mức độ tốt của lời giải:

$$\text{Fitness} = f(\text{Cá thể})$$

- Cá thể có giá trị thích nghi càng cao thì càng tốt.
- *Ví dụ:* Trong bài toán sắp xếp chỗ ngồi đám cưới, hàm thích nghi có thể đánh giá dựa trên tổng mức độ hài lòng khi ngồi cạnh nhau của các khách mời.

### 3. Lựa chọn (Selection):

- **Mục tiêu:** Chọn ra các cá thể ưu tú để làm cha mẹ cho thế hệ kế tiếp, dựa trên giá trị thích nghi.
- **Phương pháp:**
  - *Lựa chọn tỉ lệ theo độ thích nghi (Roulette Wheel Selection):* Xác suất chọn cá thể  $i$  tỉ lệ thuận với giá trị thích nghi của nó:

$$P(i) = \frac{f(i)}{\sum_{j=1}^N f(j)}$$

- *Lựa chọn giải đấu (Tournament Selection):* Chọn ngẫu nhiên một nhóm cá thể và chọn cá thể tốt nhất trong nhóm.
- *Lựa chọn theo xếp hạng (Rank Selection):* Sắp xếp quần thể theo giá trị thích nghi, rồi chọn cá thể theo thứ hạng.

### 4. Lai ghép (Crossover):

- **Mục tiêu:** Tạo ra các cá thể con mang đặc điểm di truyền từ hai cá thể cha mẹ.
- **Thực hiện:**
  - Chọn ngẫu nhiên hai cá thể cha mẹ từ quần thể dựa trên kết quả của bước lựa chọn.
  - Áp dụng các phương pháp lai ghép:

- \* *Lai một điểm (Single Point Crossover)*: Chọn một điểm ngẫu nhiên trên chuỗi gen, sau đó hoán đổi phần gen từ điểm đó trở đi giữa hai cha mẹ.
  - \* *Lai hai điểm (Two Point Crossover)*: Chọn hai điểm cắt và hoán đổi đoạn gen giữa hai điểm đó.
  - \* *Lai đồng nhất (Uniform Crossover)*: Mỗi gen của cá thể con được chọn ngẫu nhiên từ một trong hai cha mẹ.
- Xác suất lai ghép thường được thiết lập trong khoảng từ 0.7 đến 0.9.

#### 5. Đột biến (Mutation):

- **Mục tiêu:** Tăng tính đa dạng cho quần thể và tránh rơi vào cực trị địa phương.
- **Thực hiện:**
  - Xác suất đột biến thường rất thấp, từ 0.01 đến 0.1.
  - Các phương pháp đột biến:
    - \* *Đột biến nhị phân (Bit Flip)*: Đảo ngược một bit (0 thành 1 và ngược lại).
    - \* *Đột biến hoán vị (Swap Mutation)*: Hoán đổi vị trí của hai gen trong chuỗi.
    - \* *Đột biến trượt (Slide Mutation)*: Dịch chuyển một đoạn gen từ vị trí này sang vị trí khác.

#### 6. Tạo quần thể mới (Replacement):

- **Mục tiêu:** Tạo ra thế hệ mới từ thế hệ hiện tại và các cá thể con mới được tạo ra.
- **Phương pháp:**
  - *Thay thế toàn phần*: Thế hệ cũ được thay thế hoàn toàn bởi thế hệ con.
  - *Thay thế một phần*: Kết hợp cá thể cũ và con, giữ lại các cá thể có giá trị thích nghi cao nhất.

#### 7. Kiểm tra điều kiện dừng (Termination):

- **Mục tiêu:** Xác định khi nào giải thuật nên dừng lại.
- **Điều kiện dừng:**
  - Số thế hệ đạt giới hạn tối đa.
  - Không có sự cải thiện đáng kể về giá trị thích nghi sau một số thế hệ nhất định.

- Quần thể hội tụ, tức là hầu hết các cá thể đều có giá trị thích nghi tương đương nhau.

## 8. Xuất kết quả (Output):

- **Mục tiêu:** Đưa ra lời giải tối ưu hoặc cận tối ưu sau khi giải thuật kết thúc.
- **Kết quả:** Cá thể có giá trị thích nghi cao nhất trong quần thể cuối cùng sẽ được chọn làm lời giải tối ưu cho bài toán.

Giải thuật di truyền hoạt động dựa trên việc mô phỏng cơ chế chọn lọc tự nhiên, lai ghép và đột biến nhằm tìm kiếm lời giải tối ưu. Bằng cách kết hợp giữa khai phá (*exploration*) và khai thác (*exploitation*), GA có khả năng tìm kiếm trên không gian rộng lớn và vượt qua các cực trị cục bộ, mang lại hiệu quả cao trong việc giải quyết các bài toán tối ưu hóa phức tạp.

### 2.1.3. Ưu nhược điểm

Giải thuật di truyền (Genetic Algorithm – GA) là một phương pháp tối ưu hóa mạnh mẽ, đặc biệt hữu ích trong các bài toán có không gian tìm kiếm phức tạp và nhiều cực trị cục bộ. Tuy nhiên, bên cạnh những ưu điểm nổi bật, GA cũng tồn tại một số nhược điểm cần cân nhắc khi áp dụng.

#### Ưu điểm:

- *Tìm kiếm toàn cục tốt:* GA có khả năng tìm kiếm toàn cục nhờ vào sự kết hợp giữa các quá trình lai ghép và đột biến, giúp thoát khỏi bẫy của các cực trị cục bộ – vấn đề mà nhiều phương pháp tối ưu truyền thống như *gradient descent* thường gặp phải. Việc duy trì một quần thể đa dạng qua nhiều thế hệ tạo cơ hội khám phá những vùng không gian tìm kiếm tiềm năng hơn.
- *Không yêu cầu thông tin về bài toán:* Một trong những điểm mạnh đáng kể của GA là không cần bất kỳ thông tin nào về cấu trúc toán học của bài toán. Không cần tính đạo hàm, không cần liên tục, và cũng không yêu cầu bài toán phải tuyến tính. Điều này giúp GA linh hoạt trong việc giải quyết các bài toán phi tuyến tính, có nhiều cực trị cục bộ hoặc chứa nhiều ràng buộc phức tạp.
- *Khả năng xử lý nhiều ràng buộc:* GA đặc biệt phù hợp với những bài toán có nhiều ràng buộc hoặc điều kiện phức tạp. Bằng cách điều chỉnh hàm thích nghi, người thiết kế có thể hướng thuật toán tới việc tìm ra các lời giải vừa tối ưu, vừa thỏa mãn các ràng buộc của bài toán.
- *Tính linh hoạt và dễ áp dụng:* GA có thể áp dụng cho nhiều loại bài toán khác nhau như tối ưu hóa tổ hợp, tìm cực trị của hàm số, lập lịch, thiết kế hệ thống và thậm chí là trong



học máy. Chỉ cần xác định được cách mã hóa lời giải và xây dựng hàm thích nghi phù hợp, GA có thể được áp dụng dễ dàng.

- *Tính song song hóa cao:* Quá trình đánh giá độ thích nghi của từng cá thể trong quần thể có thể thực hiện song song, giúp tăng tốc độ tính toán đáng kể. Trên các hệ thống đa luồng hoặc các cụm máy tính, GA có thể khai thác tối đa tài nguyên để giảm thời gian tìm kiếm.

#### **Nhược điểm:**

- *Chi phí tính toán lớn:* Một trong những nhược điểm lớn nhất của GA là chi phí tính toán cao. Việc duy trì và đánh giá một quần thể lớn qua nhiều thế hệ đòi hỏi khối lượng tính toán rất lớn. Đặc biệt, với các bài toán phức tạp hoặc có không gian tìm kiếm rất rộng, thời gian chạy của GA có thể kéo dài đáng kể.
- *Không đảm bảo tìm được lời giải tối ưu toàn cục:* Mặc dù GA có khả năng tránh rơi vào cực trị cục bộ, nhưng không có gì đảm bảo rằng lời giải cuối cùng tìm được là tối ưu toàn cục. Trong một số trường hợp, GA có thể hội tụ sớm vào một cực trị cục bộ nếu các tham số không được điều chỉnh phù hợp hoặc quần thể không đủ đa dạng.
- *Nhạy cảm với tham số:* Hiệu quả của GA phụ thuộc rất nhiều vào việc lựa chọn các tham số như kích thước quần thể, xác suất lai ghép, xác suất đột biến và điều kiện dừng. Một bộ tham số không phù hợp có thể dẫn đến việc thuật toán hội tụ chậm, hoặc tệ hơn là không tìm được lời giải tốt.
- *Khó khăn trong thiết kế hàm thích nghi:* Hàm thích nghi đóng vai trò cực kỳ quan trọng trong GA, vì nó quyết định hướng tiến hóa của quần thể. Việc thiết kế một hàm thích nghi phù hợp không phải lúc nào cũng dễ dàng, đặc biệt đối với những bài toán có nhiều tiêu chí cần tối ưu hóa hoặc nhiều ràng buộc phức tạp. Một hàm thích nghi không tốt có thể khiến thuật toán bị dẫn dắt theo hướng sai lệch và cho ra các lời giải không mong muốn.
- *Hiệu quả giảm với bài toán đơn giản:* Đối với những bài toán có không gian tìm kiếm nhỏ hoặc có phương pháp giải chính quy, GA thường không hiệu quả bằng các phương pháp truyền thống. Trong các trường hợp này, việc sử dụng GA có thể dẫn đến lãng phí tài nguyên tính toán mà không mang lại nhiều lợi ích.

Giải thuật di truyền là một công cụ mạnh mẽ, đặc biệt phù hợp cho các bài toán tối ưu hóa phức tạp và có không gian tìm kiếm rộng lớn. Tuy nhiên, để áp dụng hiệu quả, người thiết kế cần hiểu rõ về bài toán cụ thể, lựa chọn các tham số phù hợp và thiết kế hàm thích nghi hợp lý. Việc điều chỉnh những yếu tố này sẽ giúp GA phát huy tối đa khả năng của mình và mang lại những lời giải chất lượng cao.

### 2.1.4. Ứng dụng

Giải thuật di truyền (Genetic Algorithm – GA) đã được ứng dụng rộng rãi trong nhiều lĩnh vực khác nhau nhờ khả năng tìm kiếm lời giải tối ưu hiệu quả trong các không gian phức tạp. Dưới đây là một số ứng dụng tiêu biểu của GA trong thực tế:

- **Tối ưu hóa tổ hợp:** GA được sử dụng trong các bài toán tối ưu hóa tổ hợp phức tạp như bài toán lập lịch, bài toán cái túi, và bài toán người du lịch (TSP – *Traveling Salesman Problem*). Trong các bài toán này, GA giúp tìm ra các phương án tối ưu trong việc phân bổ nguồn lực, sắp xếp công việc hoặc tối ưu hóa lộ trình.
- **Trí tuệ nhân tạo và học máy:** GA hỗ trợ huấn luyện mạng nơ-ron nhân tạo, tối ưu hóa các tham số như trọng số và cấu trúc mạng để cải thiện hiệu suất. Ngoài ra, GA còn được ứng dụng trong việc tối ưu hóa siêu tham số, giúp lựa chọn những giá trị tốt nhất nhằm nâng cao độ chính xác của mô hình học máy.
- **Xử lý dữ liệu và khai phá tri thức:** Trong lĩnh vực này, GA được áp dụng để chọn lọc các đặc trưng quan trọng trong tập dữ liệu, giảm độ phức tạp của mô hình và nâng cao hiệu quả dự đoán. Bên cạnh đó, GA cũng hỗ trợ việc phân cụm dữ liệu, giúp phát hiện các nhóm đối tượng tiềm năng trong không gian dữ liệu lớn.
- **Tối ưu hóa trong kỹ thuật và công nghiệp:** GA được ứng dụng để tối ưu hóa thiết kế mạch điện tử, giảm chi phí sản xuất và nâng cao hiệu suất hoạt động. Ngoài ra, GA còn được sử dụng trong quản lý chuỗi cung ứng, tối ưu hóa quy trình vận chuyển hàng hóa và phân bổ nguồn lực nhằm giảm chi phí và tăng hiệu quả.
- **Sinh học và y học:** Trong lĩnh vực sinh học, GA được dùng để phân tích trình tự DNA, dự đoán cấu trúc protein và tìm kiếm các mẫu gen đặc biệt. Trong y học, GA hỗ trợ quá trình phát triển thuốc bằng cách tối ưu hóa cấu trúc phân tử, nâng cao hiệu quả điều trị và giảm tác dụng phụ.
- **Ứng dụng trong đời sống:** GA có vai trò quan trọng trong quy hoạch đô thị, tối ưu hóa hệ thống giao thông, phân bổ không gian xanh và giảm thiểu tác động môi trường. Bên cạnh đó, GA còn được áp dụng trong việc thiết kế trí tuệ nhân tạo cho các NPC trong trò chơi, giúp các nhân vật này có khả năng tự học và thích nghi với hành vi của người chơi.
- **Quản lý sự kiện:** GA được áp dụng vào việc sắp xếp chỗ ngồi trong các sự kiện lớn như đám cưới, hội nghị, nhằm tối ưu hóa sự hài lòng của khách mời dựa trên các tiêu chí như sở thích, mối quan hệ hoặc tránh xung đột.

## 2.2 Các thành phần chính

Giải thuật di truyền (Genetic Algorithm – GA) mô phỏng quá trình tiến hóa tự nhiên, trong đó các cá thể “tốt” hơn sẽ có cơ hội sinh sản và truyền đặc điểm cho thế hệ sau. Để thực hiện được quá trình này, GA bao gồm một số thành phần chính như sau:

### 2.2.1. Quần thể (Population)

Quần thể là tập hợp các cá thể đại diện cho những lời giải tiềm năng của bài toán.

- **Kích thước quần thể:**

- Ảnh hưởng trực tiếp đến hiệu quả của GA.
- Quần thể lớn giúp duy trì sự đa dạng di truyền, tăng khả năng tìm ra lời giải tối ưu toàn cục. Tuy nhiên, điều này cũng làm tăng chi phí tính toán.
- Quần thể nhỏ giúp giảm thời gian tính toán nhưng có nguy cơ rơi vào cực trị cục bộ do thiếu sự đa dạng.

- **Khởi tạo quần thể:**

- Thông thường, quần thể ban đầu được tạo ngẫu nhiên để đảm bảo tính đa dạng.
- Trong một số trường hợp, quần thể ban đầu có thể được khởi tạo dựa trên những giải pháp gần đúng hoặc kinh nghiệm trước đó nhằm tăng tốc quá trình hội tụ.

### 2.2.2. Cá thể (Chromosome)

Mỗi cá thể là một lời giải tiềm năng được biểu diễn bằng một chuỗi các gen.

- **Mã hóa cá thể:**

- *Mã nhị phân:* Dùng chuỗi bit 0 và 1 để biểu diễn lời giải, phù hợp với các bài toán tối ưu tổ hợp.
- *Mã thực:* Dùng các số thực để biểu diễn, phù hợp với bài toán tối ưu liên tục.
- *Mã hoán vị:* Biểu diễn thứ tự của các phần tử, thường được sử dụng trong các bài toán hoán vị như bài toán người du lịch (TSP).

- **Vai trò:** Cá thể chứa thông tin di truyền được truyền qua các thế hệ thông qua các quá trình lai ghép và đột biến.

### 2.2.3. Gen (Gene)

Gen là đơn vị nhỏ nhất trong một cá thể, thể hiện một thuộc tính cụ thể của lời giải.

- **Ví dụ:**

- Trong bài toán xếp chỗ ngồi đám cưới, mỗi gen có thể đại diện cho vị trí của một khách mời.
- Trong bài toán tối ưu lộ trình, mỗi gen có thể biểu diễn một điểm dừng trong lộ trình đó.

- **Vai trò:** Tập hợp các gen tạo nên một cá thể, và thông tin di truyền được thể hiện thông qua việc sắp xếp hoặc giá trị của các gen.

### 2.2.4. Hàm thích nghi (Fitness Function)

Hàm thích nghi đo lường chất lượng của mỗi cá thể dựa trên mức độ phù hợp của lời giải với bài toán.

- **Cách thức hoạt động:**

- Với mỗi cá thể, hàm thích nghi trả về một giá trị gọi là “giá trị thích nghi”.
- Giá trị này càng cao thì cá thể càng phù hợp và có cơ hội lớn hơn được chọn để sinh ra thế hệ sau.

- **Ví dụ:**

- Trong bài toán sắp xếp chỗ ngồi đám cưới, hàm thích nghi có thể đo lường mức độ hài lòng của khách mời dựa trên khoảng cách giữa các khách quen biết nhau.
- Trong bài toán tối ưu lộ trình, hàm thích nghi có thể tính tổng quãng đường di chuyển và đánh giá theo tiêu chí tối thiểu hóa khoảng cách.

### 2.2.5. Lựa chọn (Selection)

Lựa chọn là quá trình chọn ra các cá thể tốt nhất từ quần thể hiện tại để làm bố mẹ cho thế hệ sau. Một số phương pháp chọn lọc phổ biến gồm:

- *Chọn lọc theo tỉ lệ thích nghi:* Xác suất chọn một cá thể tỉ lệ thuận với giá trị thích nghi của nó.
- *Chọn lọc giải đấu:* Chọn ngẫu nhiên một nhóm nhỏ các cá thể, sau đó chọn ra cá thể tốt nhất trong nhóm.

- *Chọn lọc ngẫu nhiên*: Chọn các cá thể hoàn toàn ngẫu nhiên, không xét đến giá trị thích nghi (ít được sử dụng do kém hiệu quả).

### 2.2.6. Lai ghép (Crossover)

Lai ghép là quá trình kết hợp hai cá thể bố mẹ để tạo ra một hoặc nhiều cá thể con, nhằm khai thác thông tin từ cả hai bố mẹ.

- **Phương pháp lai ghép:**

- *Lai ghép một điểm*: Chọn một điểm cắt ngẫu nhiên trên chuỗi gen và hoán đổi phần sau điểm cắt giữa hai bố mẹ.
- *Lai ghép nhiều điểm*: Chọn nhiều điểm cắt và thực hiện hoán đổi xen kẽ giữa hai cá thể.
- *Lai ghép đồng nhất*: Mỗi gen của cá thể con được chọn ngẫu nhiên từ một trong hai bố mẹ.

### 2.2.7. Đột biến (Mutation)

Đột biến là quá trình thay đổi ngẫu nhiên một hoặc nhiều gen trong cá thể nhằm tạo ra sự đa dạng, tránh việc thuật toán bị mắc kẹt ở cực trị cục bộ.

- **Phương pháp đột biến:**

- *Đảo bit*: Thay đổi ngẫu nhiên một bit từ 0 thành 1 hoặc ngược lại.
- *Hoán vị*: Hoán đổi vị trí của hai gen bất kỳ trong cá thể.
- *Biến đổi giá trị*: Thay đổi giá trị của một gen theo một quy tắc nhất định (trong mã thực).

### 2.2.8. Đánh giá (Evaluation)

Sau mỗi thế hệ, tất cả các cá thể trong quần thể đều được đánh giá bằng hàm thích nghi.

- **Vai trò:**

- Xác định mức độ phù hợp của mỗi cá thể.
- Các cá thể có giá trị thích nghi cao hơn sẽ có cơ hội lớn hơn được chọn cho quá trình lai ghép và duy trì trong quần thể.

### 2.2.9. Điều kiện dừng (Termination Condition)

Thuật toán dừng lại khi thỏa mãn một trong các điều kiện sau:

- *Số thế hệ tối đa*: Thuật toán chạy hết số thế hệ đã định trước.
- *Ngưỡng thích nghi*: Giá trị thích nghi của cá thể tốt nhất đạt ngưỡng mong muốn.
- *Độ hội tụ*: Các cá thể trong quần thể trở nên quá giống nhau, cho thấy thuật toán đã hội tụ về một lời giải duy nhất.

Mỗi thành phần trong giải thuật di truyền đóng vai trò quan trọng trong việc giúp thuật toán tìm kiếm lời giải tối ưu. Việc thiết kế cẩn thận từng thành phần sẽ ảnh hưởng trực tiếp đến hiệu suất và độ chính xác của giải pháp tìm được.

## 2.3 Ứng dụng giải thuật di truyền trong các bài toán tối ưu

Giải thuật di truyền (Genetic Algorithm – GA) được áp dụng rộng rãi trong nhiều lĩnh vực, đặc biệt là các bài toán tối ưu hóa nhờ khả năng tìm kiếm lời giải tốt trong những không gian rất lớn và phức tạp. GA không yêu cầu kiến thức sâu về bài toán mà tận dụng quá trình tiến hóa để cải thiện lời giải qua từng thế hệ. Dưới đây là các ứng dụng chi tiết:

### 2.3.1. Bài toán tối ưu tổ hợp

Bài toán tối ưu tổ hợp là những bài toán yêu cầu tìm kiếm một cách sắp xếp hoặc tổ hợp các phần tử sao cho thỏa mãn các ràng buộc nhất định và đạt giá trị tối ưu của một hàm mục tiêu.

- **Bài toán người du lịch (Traveling Salesman Problem – TSP):**
  - *Mô tả*: Tìm hành trình ngắn nhất để đi qua tất cả các thành phố đúng một lần và quay về điểm xuất phát.
  - *Ứng dụng*: GA mã hóa hành trình của người du lịch thành một chuỗi gen, trong đó mỗi gen đại diện cho một thành phố. Các phép lai ghép và đột biến được thực hiện trên các hành trình để tìm ra quãng đường tối ưu.
  - *Kết quả*: GA có thể tìm ra lộ trình gần tối ưu cho những bài toán lớn với hàng trăm thành phố, trong khi các phương pháp cổ điển thường gặp khó khăn do độ phức tạp tính toán tăng theo cấp số mũ.
- **Bài toán lập thời khóa biểu:**

- *Mô tả*: Sắp xếp lịch học cho các lớp, giảng viên và phòng học sao cho không có sự trùng lặp và tối ưu việc sử dụng tài nguyên.
- *Ứng dụng*: Mỗi cá thể biểu diễn một phương án xếp lịch, và hàm thích nghi sẽ đánh giá mức độ trùng lặp, sử dụng phòng học hợp lý và phân phối thời gian giảng dạy.
- *Kết quả*: GA tìm ra lịch học thỏa mãn tối đa các ràng buộc trong thời gian ngắn hơn nhiều so với các phương pháp vét cạn.

• **Bài toán sắp xếp chỗ ngồi đám cưới:**

- *Mô tả*: Tìm cách bố trí chỗ ngồi sao cho những người quen biết hoặc thân thiết ngồi gần nhau, giảm thiểu sự khó xử.
- *Ứng dụng*: GA mã hóa mỗi cá thể là một phương án sắp xếp khách mời. Hàm thích nghi đánh giá mức độ hài lòng dựa vào khoảng cách giữa những người quen biết.
- *Kết quả*: Giải pháp cuối cùng do GA tìm ra sẽ giảm thiểu tối đa sự bất tiện cho khách mời, giúp buổi tiệc diễn ra suôn sẻ.

### 2.3.2. Bài toán tối ưu hóa trong kỹ thuật và sản xuất

Trong lĩnh vực kỹ thuật và sản xuất, GA được áp dụng để cải thiện hiệu suất và tối ưu hóa quy trình.

• **Tối ưu hóa quy trình sản xuất:**

- *Mô tả*: Sắp xếp thứ tự công việc hoặc bố trí máy móc trong dây chuyền sản xuất nhằm giảm thiểu thời gian và chi phí.
- *Ứng dụng*: GA mã hóa thứ tự thực hiện các công đoạn trong quy trình sản xuất thành một cá thể. Các phép lai ghép và đột biến sẽ hoán đổi vị trí các công đoạn để tìm ra thứ tự tối ưu.
- *Kết quả*: Giảm thời gian sản xuất, tăng hiệu suất vận hành và giảm lãng phí nguyên vật liệu.

• **Thiết kế mạng lưới giao thông:**

- *Mô tả*: Tối ưu hóa vị trí đặt trạm thu phí, đèn giao thông hoặc tìm lộ trình di chuyển tối ưu.
- *Ứng dụng*: GA giúp tìm ra cấu hình tối ưu của mạng lưới giao thông bằng cách đánh giá thời gian di chuyển và mức độ ùn tắc của từng phương án.

- *Kết quả*: Tối ưu hóa thời gian di chuyển, giảm thiểu tình trạng kẹt xe.

### 2.3.3. Tối ưu hóa trong trí tuệ nhân tạo và học máy

GA cũng được ứng dụng trong việc huấn luyện và cải thiện hiệu quả của các mô hình trí tuệ nhân tạo (AI) và học máy (Machine Learning).

- **Tối ưu tham số của mạng nơ-ron nhân tạo:**

- *Mô tả*: Tìm kiếm cấu trúc mạng nơ-ron tối ưu, bao gồm số lượng lớp ẩn, số nơ-ron trong mỗi lớp, và các siêu tham số như tốc độ học (*learning rate*) hoặc hệ số phạt (*regularization*).
- *Ứng dụng*: Mỗi cá thể trong quần thể là một cấu hình mạng nơ-ron khác nhau. Hàm thích nghi đánh giá độ chính xác trên tập kiểm tra, và GA sẽ tiến hóa dần dần để tìm ra cấu hình tối ưu.
- *Kết quả*: Cải thiện đáng kể hiệu suất của mô hình mà không cần thử nghiệm thủ công từng tham số.

- **Chọn đặc trưng (Feature Selection):**

- *Mô tả*: Lựa chọn các đặc trưng quan trọng nhất trong tập dữ liệu nhằm giảm bớt độ phức tạp và tăng độ chính xác của mô hình.
- *Ứng dụng*: GA mã hóa mỗi cá thể là một tập hợp các đặc trưng được chọn. Hàm thích nghi đánh giá mô hình dựa trên độ chính xác khi huấn luyện với tập đặc trưng đó.
- *Kết quả*: Giảm thời gian huấn luyện, tăng độ chính xác và tránh hiện tượng *overfitting*.

### 2.3.4. Bài toán tối ưu tài chính và kinh tế

Trong lĩnh vực tài chính và kinh tế, GA hỗ trợ ra quyết định và tối ưu hóa lợi nhuận.

- **Tối ưu hóa danh mục đầu tư:**

- *Mô tả*: Xây dựng danh mục đầu tư tối ưu với mức rủi ro thấp nhất và lợi nhuận cao nhất.
- *Ứng dụng*: GA mã hóa mỗi cá thể là một danh mục đầu tư, với tỉ trọng của từng loại tài sản là các gen. Hàm thích nghi sẽ đánh giá lợi nhuận kỳ vọng và rủi ro của mỗi danh mục.
- *Kết quả*: Giúp nhà đầu tư đưa ra các quyết định hợp lý và giảm thiểu rủi ro.



- **Dự đoán thị trường:**

- *Mô tả:* Xây dựng mô hình dự đoán biến động giá cổ phiếu hoặc tỷ giá hối đoái.
- *Ứng dụng:* GA tìm ra bộ tham số tối ưu cho các mô hình dự đoán như mạng nơ-ron hoặc cây quyết định.
- *Kết quả:* Cải thiện độ chính xác trong dự đoán, hỗ trợ ra quyết định đầu tư.

### 2.3.5. Tối ưu hóa trong đời sống hàng ngày

GA còn được áp dụng vào những vấn đề thực tiễn trong cuộc sống:

- **Lập kế hoạch du lịch:** Tìm ra hành trình tối ưu dựa trên chi phí, thời gian, và sở thích cá nhân.
- **Quản lý kho hàng:** Sắp xếp hàng hóa trong kho sao cho thời gian và chi phí vận chuyển được giảm thiểu.

Giải thuật di truyền đã chứng tỏ tính hiệu quả trong việc giải quyết các bài toán tối ưu phức tạp, đặc biệt là những bài toán có không gian nghiệm lớn. Việc mô phỏng quá trình tiến hóa tự nhiên giúp GA tìm kiếm lời giải tiềm năng một cách nhanh chóng và linh hoạt. Khả năng áp dụng của GA rất rộng, từ tối ưu hóa quy trình sản xuất, quản lý tài chính đến việc giải quyết các vấn đề trong cuộc sống hàng ngày.

## Chương 3

# THIẾT KẾ VÀ GIẢI THUẬT

### 3.1 Mô hình bài toán

#### Xác định bài toán:

Bài toán yêu cầu sắp xếp một danh sách khách mời vào các bàn tiệc sao cho tổng mức độ thân thiết giữa các khách mời ngồi cùng bàn là cao nhất, đồng thời tuân thủ các ràng buộc thực tế.

#### Đầu vào:

- Danh sách khách mời (ví dụ:  $N$  khách).
- Ma trận điểm thân thiết giữa các cặp khách mời (dựa trên bảng điểm: vợ/chồng: 2000, anh/chị/em: 900, bạn bè: 100, v.v.).
- Số bàn tiệc ( $M$ ) và số người tối đa trên mỗi bàn ( $K$ ).

#### Đầu ra:

Một sơ đồ sắp xếp chỗ ngồi, trong đó mỗi khách mời được gán vào một bàn, sao cho tổng điểm thân thiết toàn cục được tối đa hóa và các ràng buộc được thỏa mãn.

#### Mô hình hóa

- Xem mỗi khách mời là một “đối tượng” cần được phân bổ vào  $M$  bàn.
- Mỗi bàn là một tập hợp con của khách mời, với kích thước không vượt quá  $K$ .
- Mức tiêu chuẩn tối ưu hóa:

$$\max \sum_{\text{mỗi bàn}} \sum_{\text{các cặp trong bàn}} C(i, j),$$

trong đó  $C(i, j)$  là điểm thân thiết giữa khách  $i$  và  $j$  trong cùng bàn.

- Ràng buộc: số khách trong mỗi bàn  $\leq K$ , và các yêu cầu đặc biệt (nếu có).

**Ví dụ minh họa:**

Với 6 khách (A, B, C, D, E, F), 2 bàn, mỗi bàn tối đa 3 người, và ma trận điểm thân thiết đã cho (A-B: 2000, C-D: 900, v.v.), mục tiêu là tìm cách gán sao cho tổng điểm cao nhất.

## 3.2 Mã hóa cá thể

**Ý nghĩa** Mỗi cá thể trong quần thể GA đại diện cho một giải pháp khả thi – tức là một cách sắp xếp chỗ ngồi. Cách mã hóa phải đảm bảo tính hợp lệ và dễ áp dụng các toán tử di truyền.

**Phương pháp mã hóa**

### Mã hóa đầy đủ trên hoán vị

- Cá thể là một dãy hoán vị của  $N$  khách mời.
- Chia dãy thành  $M$  đoạn, mỗi đoạn tương ứng với một bàn (đoạn có độ dài tối đa  $K$ ).
- *Ví dụ:* Với 6 khách, cá thể  $[3, 1, 5, 2, 4, 6]$  có thể được chia thành: Bàn 1: C, A, E (vị trí 1-3); Bàn 2: B, D, F (vị trí 4-6).
- *Ưu điểm:* Đơn giản, đảm bảo mỗi khách chỉ thuộc một bàn.
- *Nhược điểm:* Có thể cần điều chỉnh nếu số khách không chia đều cho số bàn.

### Mã hóa nhị phân

- Dựng ma trận  $N \times M$ , trong đó  $a_{i,j} = 1$  nếu khách  $i$  ngồi bàn  $j$ , 0 nếu không.
- *Ràng buộc:*
  - Mỗi hàng chỉ có duy nhất một giá trị 1 (mỗi khách chỉ ngồi một bàn).
  - Mỗi cột có tối đa  $K$  giá trị 1 (giới hạn số người mỗi bàn).
- *Ví dụ:* Với 6 khách, 2 bàn: Cá thể có thể là  $[[1, 0], [0, 1], [1, 0], [0, 1], [1, 0], [0, 1]] \rightarrow$  Bàn 1: A, C, E; Bàn 2: B, D, F.
- *Ưu điểm:* Linh hoạt với ràng buộc bổ sung.
- *Nhược điểm:* Phức tạp hơn khi thực thi.

**Lựa chọn** Trong bài toán này, mã hóa hoán vị thường được chọn vì đơn giản, dễ áp dụng toán tử lai ghép và đột biến, đồng thời phù hợp với bài toán phân bố tuần tự.

### 3.3 Hàm thích nghi (Fitness Function)

**Vai trò** Hàm thích nghi đánh giá chất lượng của mỗi cá thể, định hướng quá trình tiến hóa của GA.

**Công thức** Tính tổng điểm thân thiết của tất cả cặp khách mời trong cùng bàn:

$$\text{Fitness} = \sum_{\text{mỗi bàn } i, j \text{ trong bàn}} \sum_{i < j} C(i, j)$$

Trong đó  $C(i, j)$  là điểm thân thiết giữa khách  $i$  và  $j$ .

Nếu có vi phạm ràng buộc:

- Qua số người tối đa mỗi bàn: Trừ một hình phạt lớn, ví dụ:  $\text{Fitness} = \text{Fitness} - 10000 \times \text{số người vượt quá}$ .
- Không thỏa mãn yêu cầu đặc biệt (ví dụ: hai người không được ngồi cùng): Trừ thêm hình phạt, ví dụ:  $-5000$  cho mỗi vi phạm.

**Ví dụ**

- Cá thể: Bàn 1: A, B, C; Bàn 2: D, E, F.
- Bàn 1: A-B: 2000, A-C: 700, B-C: 0  $\rightarrow$  Tổng: 2700.
- Bàn 2: D-E: 0, D-F: 0, E-F: 100  $\rightarrow$  Tổng: 100.
- $\text{Fitness} = 2700 + 100 = 2800$  (nếu không vi phạm ràng buộc).

### 3.4 Toán tử lai ghép và đột biến

**Mục đích** Kết hợp hai cá thể cha mẹ để tạo ra cá thể con, kế thừa đặc điểm tốt.

**Phương pháp Lai ghép theo thứ tự (Order Crossover):**

- Chọn ngẫu nhiên một đoạn từ cha mẹ 1, sao chép vào con.
- Điền phần còn lại theo thứ tự từ cha mẹ 2, tránh lặp.
- *Ví dụ:* Cha mẹ 1: [1, 2, 3, 4, 5, 6], Cha mẹ 2: [4, 5, 2, 1, 6, 3]. Chọn đoạn [2, 3] từ Cha mẹ 1  $\rightarrow$  Con: [x, x, 2, 3, x, x], điền từ Cha mẹ 2: [4, 5, 2, 3, 1, 6].

Tỷ lệ lai ghép: Thường chọn 0.7-0.9 (70-90% cá thể được lai ghép).

Ý nghĩa Giữ cấu trúc tốt từ cha mẹ, tạo ra sự kết hợp mới.

### 3.4.1 Toán tử đột biến (Mutation)

Mục đích Tăng tính đa dạng, tránh hội tụ sớm về cực trị địa phương.

### 3.4.2 Phương pháp

**Hoán đổi (Swap Mutation):**

- Chọn ngẫu nhiên hai vị trí trong cá thể, hoán đổi khách mời.
- Ví dụ:  $[1, 2, 3, 4, 5, 6] \rightarrow [1, 5, 3, 4, 2, 6]$  (hoán đổi vị trí 2 và 5).

Tỷ lệ đột biến: Thường thấp, khoảng 0.01-0.05 (1-5% cá thể bị đột biến).

Ý nghĩa Đảm bảo khám phá các giải pháp mới, đặc biệt khi quần thể bắt đầu giống nhau.

## 3.5 Chiến lược chọn lọc

### 3.5.1 Mục đích

Chiến lược chọn lọc nhằm chọn ra các cá thể tốt nhất từ quần thể hiện tại để làm cha mẹ cho thế hệ sau, đảm bảo rằng các đặc điểm tốt được giữ lại và cải thiện qua các thế hệ.

### 3.5.2 Phương pháp

**Lựa chọn giải đấu (Tournament Selection):**

- Chọn ngẫu nhiên  $k$  cá thể từ quần thể (thường  $k = 2$  hoặc  $3$ ).
- So sánh giá trị thích nghi của các cá thể trong nhóm, chọn cá thể có giá trị thích nghi cao nhất làm cha mẹ.
- Ví dụ: Với  $k = 2$ , chọn 2 cá thể ngẫu nhiên: Cá thể 1 (Fitness = 2800), Cá thể 2 (Fitness = 2500)  $\rightarrow$  Chọn Cá thể 1.

**Lựa chọn giải đấu (Tournament Selection):**

- Giữ lại một số cá thể tốt nhất (ví dụ: 1-2 cá thể) từ thế hệ cũ, đảm bảo không mất giải pháp tốt.
- Kết hợp với các cá thể con từ lai ghép và đột biến để tạo quần thể mới.

### 3.5.3 Quy trình:

- Quần thể ban đầu: 100 cá thể.
- Sau mỗi thế hệ: Giữ 2 cá thể tốt nhất (elitism), chọn 98 cá thể còn lại bằng giải đấu từ tập hợp cha mẹ và con c

# Chương 4

## Kết Quả

### 4.1 Thiết lập dữ liệu

#### 4.1.1 Môi trường thực hiện

- **Ngôn ngữ lập trình:** Python 3.x.
- **Thư viện hỗ trợ:** pandas để đọc dữ liệu từ file Excel, random để xử lý ngẫu nhiên trong GA.
- **Phần cứng:** Máy tính cá nhân.

#### 4.1.2 Dữ liệu đầu vào

- **Danh sách khách mời:** 50 khách (Khach1 đến Khach50), lấy từ file Excel `data.xlsx`.
- **Quan hệ:** Ma trận quan hệ giá trị các cặp khách mời, với điểm thân thiết:
  - Vợ/chồng: 2000 điểm.
  - Anh/chị/em ruột: 900 điểm.
  - Cha/mẹ-con: 700 điểm.
  - Anh/chị/em họ: 500 điểm.
  - Họ hàng (dì/chú/bác-cháu): 300 điểm.
  - Bạn bè: 100 điểm.
  - Không quen: 0 điểm.

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS COMMENTS
PS D:\Vjet_Y2\KY_2\AI\BTL> & C:/Users/tengu/AppData/Local/Programs/Python/Python312/python.exe d:\Vjet_Y2\KY_2\AI\BTL\demo.py
Sơ đồ chỗ ngồi tối ưu:
Bàn 1: Khách 4, Khách 2, Khách 44, Khách 4
Bàn 2: Khách 3, Khách 1, Khách 7, Khách 7
Bàn 3: Khách 2, Khách 2, Khách 17, Khách 1
Bàn 4: Khách 28, Khách 3, Khách 1, Khách 28
Bàn 5: Khách 1, Khách 28, Khách 1, Khách 44
Bàn 6: Khách 22, Khách 4, Khách 2, Khách 4
Bàn 7: Khách 1, Khách 44, Khách 4, Khách 2
Bàn 8: Khách 2, Khách 2, Khách 19, Khách 4
Bàn 9: Khách 2, Khách 2, Khách 4, Khách 28
Bàn 10: Khách 1, Khách 1, Khách 44, Khách 2
Bàn 11: Khách 28, Khách 1, Khách 44, Khách 1
Bàn 12: Khách 3, Khách 13, Khách 2, Khách 13
Bàn 13: Khách 48, Khách 43
Điểm cao nhất: 102900

```

Hình 4.1: Kết quả chạy được

### 4.1.3 Ràng buộc

- Số người tối đa mỗi bàn: 4 người (`gioi_han_so_nguoi = 4`).
- Tổng số bàn: 13 bàn (50 khách chia 4, dư 2 khách được gán vào bàn cuối).

#### File dữ liệu

- D:\Vjet\_Y2\KY\_2\AI\BTL\data.xlsx, chứa danh sách quan hệ như trong tài liệu.

### 4.1.4 Thông số GA

- **Kích thước quần thể** (`KICH_THUOC_QUAN_THE`): 100 cá thể.
- **Số thế hệ tối đa** (`SO_THE_HE`): 1000 thế hệ.
- **Tỷ lệ đột biến** (`TY_LE_DOT_BIEN`): 0.1 (10%).
- **Toán tử lai ghép**: Lai ghép ngẫu nhiên chọn từng bàn từ cha hoặc mẹ (`lai_ghep`).
- **Toán tử đột biến**: Hoán đổi ngẫu nhiên vị trí hai khách trong danh sách phân (`dot_bien`).
- **Chiến lược chọn lọc**: Chọn 50% cá thể tốt nhất từ quần thể (`chon_loc`).

### 4.1.5 Phương pháp chạy

- Thuật toán được chạy 5 lần để lấy trung bình fitness và giảm ảnh hưởng của yếu tố ngẫu nhiên.
- Thời gian chạy được ghi nhận để đánh giá hiệu suất.



## 4.2 Kết quả thu được

## 4.3 Đánh giá hiệu suất của giải thuật

Với 50 khách và không gian tìm kiếm rất lớn (số cách chia  $\sim 10^{30}$ ), GA không thể đảm bảo tối ưu tuyệt đối, nhưng fitness trung bình 40,500 cho thấy giải pháp tốt, ưu tiên các mối quan hệ thân thiết (vợ/chồng, anh em ruột). So với phân bố ngẫu nhiên (fitness trung bình ước tính  $\sim 10,000 - 15,000$ ), GA vượt trội rõ rệt.

### 4.3.1 Khả năng đáp ứng ràng buộc

- Ràng buộc 4 người mỗi bàn được thỏa mãn hoàn toàn cách chia danh sách trong `tao_ca_the`.
- Tuy nhiên, một số chưa xử lý ràng buộc bổ sung (ví dụ: không muốn ngồi cùng), cần cải tiến hàm fitness bằng cách thêm hình phạt.

### 4.3.2 Thời gian thực thi

- Trung bình 30 giây cho 1000 thể hệ, phù hợp với bài toán tiệc cưới 50 khách.
- Với quy mô lớn hơn (100 khách), thời gian có thể tăng lên vài phút, vẫn chấp nhận được so với tìm kiếm vét cạn (gần như bất khả thi).

### 4.3.3 Ưu điểm

- Tự động tạo sơ đồ tối ưu mà không cần can thiệp thủ công.
- Linh hoạt với số khách và giới hạn bàn (cần điều chỉnh `gioi_han_so_nguoi`).
- Fitness cao, ưu tiên các mối quan hệ quan trọng (vợ/chồng, gia đình).

### 4.3.4 Hạn chế

- Tỷ lệ đột biến 0.1 có thể cao, dẫn đến mất giải pháp tốt trong một số lần chạy (thường tỷ lệ chỉ 0.01-0.05).
- Lai ghép chọn ngẫu nhiên từng bàn (`lai_ghep`) không tối ưu bằng các phương pháp như lai ghép theo thứ tự (order crossover), có thể làm giảm tính kế thừa.
- Chưa xử lý trường hợp số khách không chia đều cho số người mỗi bàn một cách tối ưu (bàn cuối ít người hơn).
- Không có elitism (giữ cá thể tốt nhất), có thể làm mất giải pháp tối ưu qua các thế hệ.

#### 4.3.5 Đề xuất cải tiến

- Giảm tỷ lệ đột biến xuống 0.02-0.05 để giữ ổn định quần thể.
- Thêm elitism, giữ 1-2 cá thể tốt nhất mỗi thế hệ.
- Sử dụng lai ghép theo thứ tự thay vì chọn ngẫu nhiên từng bàn.
- Thêm hình phạt vào hàm fitness để xử lý ràng buộc bổ sung (ví dụ: -5000 nếu hai người không muốn ngồi cùng bị gán cùng bàn).
- Tối ưu thuật toán bằng cách khởi tạo quần thể ban đầu với heuristic (nhóm các cặp vợ/chồng trước).

# Chương 5

## Kết Luận

Sau quá trình nghiên cứu và thực nghiệm, đề tài "Ứng dụng giải thuật di truyền tối ưu sắp xếp chỗ ngồi tiệc cưới" đã đạt được những kết quả đáng khích lệ. Giải thuật di truyền đã chứng tỏ tính hiệu quả trong việc tìm kiếm phương án sắp xếp tối ưu, đảm bảo sự hài hòa trong quan hệ giữa các khách mời, hạn chế tối đa xung đột và tạo ra trải nghiệm tốt nhất cho tất cả mọi người tham dự.

Thông qua đề tài này, nhóm chúng em đã có cơ hội áp dụng kiến thức về trí tuệ nhân tạo, thuật toán tối ưu hóa vào một bài toán thực tế, đồng thời trau dồi kỹ năng làm việc nhóm, tư duy giải quyết vấn đề và lập trình thuật toán. Mặc dù đã đạt được một số thành công nhất định, nhưng đề tài vẫn còn những hạn chế và có thể được cải tiến trong tương lai, chẳng hạn như mở rộng phạm vi ứng dụng cho các sự kiện khác, cải thiện tốc độ tính toán hoặc kết hợp với các công nghệ hiện đại hơn như học máy để nâng cao độ chính xác.

Chúng em hy vọng rằng nghiên cứu này sẽ là một tài liệu tham khảo hữu ích cho những ai quan tâm đến lĩnh vực tối ưu hóa và ứng dụng thuật toán vào thực tế. Một lần nữa, nhóm xin chân thành cảm ơn cô Lê Thị Thùy Trang cùng tất cả những người đã hỗ trợ và đồng hành trong suốt quá trình thực hiện đề tài này.

### 5.1 Điểm mạnh của nghiên cứu

#### 5.1.1 Tính hiệu quả của giải thuật di truyền

Giải thuật di truyền (Genetic Algorithm - GA) là một phương pháp tối ưu hóa mạnh mẽ, có khả năng tìm kiếm nghiệm tốt nhất trong không gian rất lớn mà các phương pháp truyền thống như tìm kiếm vét cạn hay tham lam khó có thể thực hiện hiệu quả. Với khả năng kết hợp đột biến, lai ghép và chọn lọc, giải thuật này giúp tạo ra các phương án sắp xếp chỗ ngồi có độ tối ưu cao, hạn chế tối đa sự không hài lòng của khách mời.

### **5.1.2 Tính linh hoạt và khả năng mở rộng**

Mô hình sắp xếp chỗ ngồi có thể dễ dàng điều chỉnh để áp dụng cho các sự kiện khác như hội nghị, tiệc công ty, sự kiện văn hóa, lễ hội... Có thể tích hợp thêm các tiêu chí khác như sở thích cá nhân, độ thân thiết giữa các khách mời, hoặc thậm chí yếu tố cá nhân hóa dựa trên dữ liệu lịch sử để nâng cao trải nghiệm của từng người tham dự.

### **5.1.3 Ứng dụng công nghệ vào đời sống thực tế**

Đề tài không chỉ mang ý nghĩa về mặt lý thuyết mà còn có tính thực tiễn cao, giúp hỗ trợ công tác tổ chức sự kiện chuyên nghiệp hơn. Các nền tảng đặt chỗ, dịch vụ tiệc cưới có thể tích hợp thuật toán này để nâng cao trải nghiệm khách hàng, giúp quá trình lên kế hoạch trở nên tự động và hiệu quả hơn.

## **5.2 Hạn chế và thách thức**

### **5.2.1 Độ phức tạp tính toán và thời gian xử lý**

Mặc dù giải thuật di truyền có khả năng tối ưu hóa tốt, nhưng khi số lượng khách mời tăng lên đáng kể (hàng trăm hoặc hàng nghìn người), thời gian xử lý có thể kéo dài do số lượng tổ hợp có thể xảy ra là rất lớn. Điều này đòi hỏi tối ưu hóa hiệu suất của thuật toán, có thể bằng cách điều chỉnh các tham số như tỉ lệ lai ghép, tỉ lệ đột biến hoặc kết hợp với các phương pháp khác như lập trình động để tăng tốc độ tìm kiếm nghiệm.

### **5.2.2 Khả năng đảm bảo nghiệm tối ưu tuyệt đối**

Giải thuật di truyền không đảm bảo tìm được nghiệm tối ưu toàn cục mà chỉ có thể tìm ra nghiệm gần đúng tốt nhất trong giới hạn thời gian và số thế hệ nhất định. Một số trường hợp đặc biệt có thể cần đến sự can thiệp của con người để điều chỉnh phương án cuối cùng, đặc biệt khi có các ràng buộc phức tạp mà thuật toán khó mô hình hóa đầy đủ.

## **5.3 Tính chủ quan trong đánh giá mức độ hài lòng**

Các tiêu chí đánh giá sự hài lòng của khách mời có thể khác nhau tùy vào từng sự kiện và từng người, điều này đặt ra thách thức trong việc xây dựng hàm đánh giá (fitness function) sao cho phản ánh chính xác nhất nhu cầu thực tế. Việc thu thập dữ liệu từ người dùng để cải thiện mô hình cũng có thể gặp khó khăn về mặt bảo mật và quyền riêng tư.

### 5.3.1 Tính ứng dụng thực tế và hướng phát triển tương lai

Với những ưu điểm đã phân tích, đề tài này có thể được ứng dụng rộng rãi trong nhiều lĩnh vực khác nhau như:

- **Tổ chức tiệc cưới, hội nghị, sự kiện quy mô lớn:** Giúp ban tổ chức dễ dàng sắp xếp chỗ ngồi hợp lý, nâng cao sự hài lòng của khách mời.
- **Hệ thống đặt chỗ thông minh:** Các nhà hàng, khách sạn có thể áp dụng thuật toán này vào hệ thống đặt bàn tự động, tối ưu hóa việc phục vụ khách hàng.
- **Quản lý chỗ ngồi trong giao thông công cộng, rạp chiếu phim:** Hỗ trợ tối ưu hóa sơ đồ chỗ ngồi dựa trên nhu cầu, thói quen hành khách.

Trong tương lai, hướng phát triển của đề tài có thể bao gồm:

- **Cải thiện thuật toán để nâng cao hiệu suất:** Kết hợp giải thuật di truyền với các phương pháp tối ưu khác như lập trình tuyến tính, học sâu để tăng tốc độ tính toán và đảm bảo độ chính xác cao hơn.